AD_____

Award Number: DAMD17-99-1-9091

TITLE: Diagnosing Breast Cancer Using Protease Fingerprint

PRINCIPAL INVESTIGATOR: Emily I. Chen

CONTRACTING ORGANIZATION: The Burnham Institute
La Jolla, California 92037

REPORT DATE: June 2002

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 2002	Annual Summary (1 Jun 01 - 31 May 02)

4. TITLE AND SUBTITLE
Diagnosing Breast Cancer Using Protease Fingerprint

5. FUNDING NUMBERS
DAMD17-99-1-9091

6. AUTHOR(S)
Emily I. Chen

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

The Burnham Institute
La Jolla, California 92037

E-Mail: echen@burnham.org

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Report contains color.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

The utility of using substrate phage display library to study disease-relevant metalloproteinase was further proved by identifying the selective substrates for MMP-14 (MT1-MMP). In addition, a different proteomic approach of achieving the original aim of identifying new diagnostic marker for breast cancer was applied and reported in this report. An activity-based probe, FP-biotin, was used to analyze the global activity pattern of a class of disease-relevant enzymes, serine hydrolases, in normal epithelial cells and several breast cancer cells. A novel potential tumor suppressor protein,human carboxylesterase-2, was identified as the result of the comparison. Initial confirmation of the cell line comparison was done by analyzing the regulation at the transcription and translation level. This result is a good demonstration of the usefulness of applying proteomic approaches to analyze complex biological samples in health and disease states.

14. SUBJECT TERMS	15. NUMBER OF PAGES
breast cancer, protease fingerprint, diagnostic marker, disease-relevant metalloproteinase, FP-biotin, enzymes	43
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

___ Where copyrighted material is quoted, permission has been obtained to use such material.

___ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

___ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

__ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

N/A In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

N/A In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

N/A In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

6/12/02

PI - Signature Date

Table of Contents

Introduction:

The purpose of my study is to obtain a collective knowledge of proteases which activity is differentially overexpressed during breast cancer progression using phage display technology. The profile of proteases or "protease fingerprint" will then be validated as a diagnostic tool for predicting breast cancer progression.

Body:

Proteases are key regulators of a wide range of physiological processes (1,2), and are recognized as important and tractable drug targets. The activity of proteases is tightly regulated by inhibitors, and some of them by protein activation. Many proteases such as matrix-metalloproteinase are synthesized as zymogen (inactive form) and are activated later by other proteases (3-5). Upon activation, protease activity is then regulated by either the non-specific inhibitors such as α1-macroglobulin or specific inhibitors such as tissue specific metalloproteinase inhibitors (TIMPs) for MMPs (6). Mistakes in the regulation of protease activity can lead to various pathological conditions such as cancer growth, tumor metastasis, inflammation, cardiovascular, and autoimmune diseases (7-11). Consequently, information of protease activity in diseases and progression of diseases will be important in discovering biomarkers and potential therapeutic targets. In this fellowship proposal, I hypothesized that a profile of protease activity can be used as a diagnostic tool to predict breast cancer progression. Three aims were designed in this fellowship proposal to test this hypothesis. So far, studies toward completing these three aims have lead to four publications. Three publications were included in last year's annual summary report. The manuscript published this year is included in this annual summary repot.

Aim#1. Construct A Substrate Phage Display Library To Use As A Tool For Studying Protease Fingerprint.

I completed this aim last year and had three publications of the method and studies of matrix metalloproteinase-2 and matrix metalloproteinase-9. As one of the future plans listed in last year's annual summary report, I devoted part of my effort on characterizing the substrate recognition profile of another breast cancer relevant protease, MT1-MMP (MMP-14) and obtaining highly selective peptide substrates.

Membrane Type-1 Matrix Metalloproteinase (MT1-MMP, MMP-14) is the prototypic member of the MT-MMPs, and its expression has been associated with a variety of cellular and developmental processes as well as multiple pathophysiologic conditions. Similar to most other MMPs, MT1-MMP cleaves a number of matrix proteins including collagen, fibronectin and vitronectin (12-14). Other work shows that MT1-MMP has substrates that extend beyond extracellular matrix proteins. MT1-MMP is known to convert pro-MMP-2 to active protease (15-18), apparently through its tri-molecular complex with TIMP-2 and the $\alpha_v\beta_3$ integrin (19,20). Recent work shows that cleavage of cell surface molecules such as CD44, pro-α_v integrin and transglutatminase by MT1-MMP modulates cell migration (21-23). The importance of the broad substrate recognition specificity of MT1-MMP is consistent with the complex phenotype of the MT1-MMP-deficient mice. These mice exhibit overt dwarfism, but close inspection also showed several additional defects including arthritis, delayed ossification of bone, and the inability to respond to angiogenic stimuli (24,25). Clearly, MT1-MMP is likely to have multiple physiologic and pathophysiologic substrate. Knowing the identity of these substrates, and the structural basis for their recognition, would provide an additional level of understanding of MT1-MMP.

To this end, we used an unbiased substrate phage selection to identify optimal peptide substrates for MT1-MMP. Four groups of substrates, with varying degrees of selectivity were identified. As anticipated, non-selective substrates comprised primarily of the Pro-X-X-$\downarrow$-X_{Hy} sequence were identified. These substrates are collagen-like, and have emerged as a canonical, and generally non-selective, MMP recognition motif (26-28). We also identified substrates that are recognized by MT1-MMP and both gelatinases (MMP-2 and MMP-9). Most significantly, a group of highly selective substrates for MT1-MMP were identified. Each of the highly specific substrates is devoid of a Pro residue at the P_3 position and contains a critical Arg at the P_4 position. Modeling of the selective and non-selective substrates bound to the catalytic pocket of MT1-MMP indicates two separate binding modes. These modes are distinguished by the degree of distortion at the P_3 position of substrate. These observations provide a solid structural basis for the recognition of distinct classes of substrates by MT1-MMP and for the design of highly specific inhibitors of this enzyme.

The results of this study are in the manuscript included in this annual summary report, and the manuscript has been accepted by J Bio Chem in April, 2002.

Aim#2. Charcterize profile of protease activity present in the blood of mice bearing non-metastatic breast cancer tumors and mice bearing metastatic breast cancer tumors.

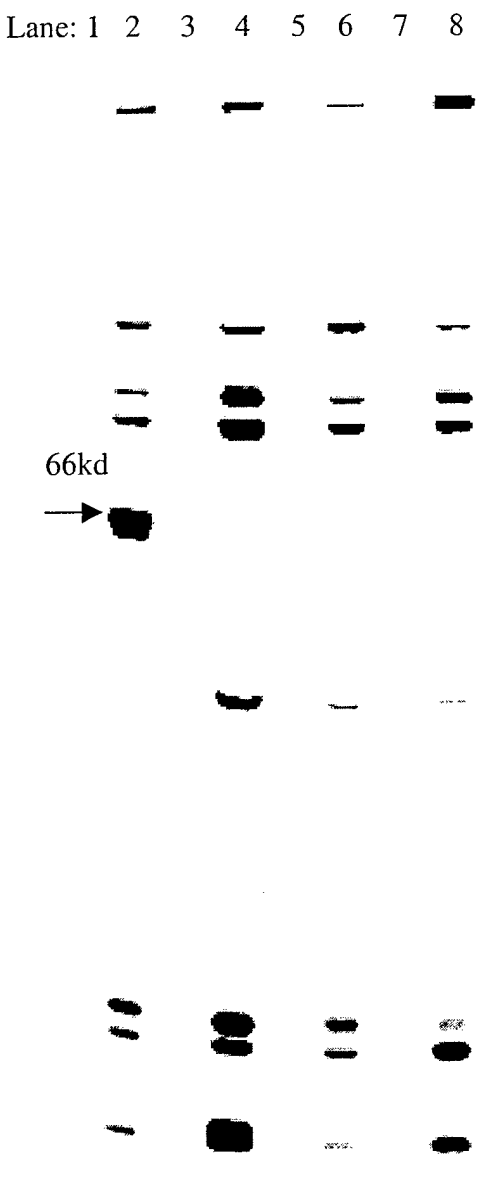
In the annual summary report last year, I reported my unsuccessful attempt to profile protease activity from breast cancer tumor using substrate phage display library. To accomplish similar goal to the study that I proposed in aim#2, I collaborated with Dr. Cravatt to use an activity-based proteomic probe, biotinylated fluorophosphonate (FP-biotin), developed in his laboratory to target the serine hydrolase family of enzyme (29). As a result of a shared catalytic mechanism, nearly all serine hydrolases are irreversibly inhibited by FP-based reagents. Additionally, the reaction of FPs with serine hydrolases depends on these enzyems being catalytically active. Thus, in principle, FP-biotin could provide a complete profile of the active serine hydrolases present in a given proteome.

Serine hydrolases represent one of the largest and most diverse families of enzymes in higher eukaryotes. Subclasses of serine hydrolases include proteases, lipases, esterases, amidases, and transacylases, with individual enzymes implicated in numerous (patho)physiological processes, such as blood coagulation (30), T cell cytotoxicity (31), neural plasticity (32), neurotransmitter catabolism (33), peptide/protein processing (34), protein/lipid digestion (35), angiogenesis (36), emphysema (37), and cancer (38). Some of the known serine proteases that play important roles in breast cancer development and metastasis such as tPA and uPA also belong to the family of serine hydrolases. Importantly, the activities of many serine hydrolases are regluated in a posttranslational manner (e.g., zymogen cleavage for catalytic activation (39), inhibitor binding for catalytic inactivation (40,41)), indicating that activity-based proteomic approach will be useful to perform functional analysis of these enzymes.

The goal of utilizing this new approach is similar to the goal of aim#2 that I proposed. I hypothesize that metastastatic related serine hydrolase(s) can be identified by comparing the serine hydrolase profile of non-metastatic and metastatic tumors. To demonstrate the usefulness of this approach, I first compared the serine hydrolase profile of normal epithelial cells (HMEC) to several breast cancer cell lines.

Result#1: A 66kd protein was down regulated in breast cancer cells.

I compared the serine hydrolase profile of HMEC (normal epithelial cells) and three breast cancer cell lines: MDA-MB-435 (highly metastatic), MDA-MB-231 (metastatic), and MCF-7 (non-metastatic). Cells from each cell line were grown under the same condition and harvested when reached the same confluence. Cell pellets from each cell line were homogenized and separated into soluble and insoluble fractions by centrifugation. Protein concentration of each fraction was measured by BCA assay (BioRad, CA) and adjusted to 1mg/ml before adding the serine hydrolase FP probe.



5uM of serine hydrolase FP probe were incubated at room temperature for 1 hour with 40ug of protein from each cell lysate. As a control for non-specific background binding, one sample from each cell lysate was heated at 90°C for 10 minutes to denature the protein before the addition of the FP probe. After 1hr of incubation, each probed sample was heated with the protein loading buffer and separated on 10% SDS-PAGE. In figure1, soluble cell lysates from each cell line were probed by FP-Tamara, a red fluorescent tagged version of FP-biotin, and revealed the serine hydrolase profile of each cell line by scanning the fluorescence of the serine hydrolase-probe complex on the gel. One band around 66kd in the soluble fraction comparison among all four cell lines appeared significantly down-regulated in breast cancer cells but abundant in normal epithelial cells.

Figure 1.

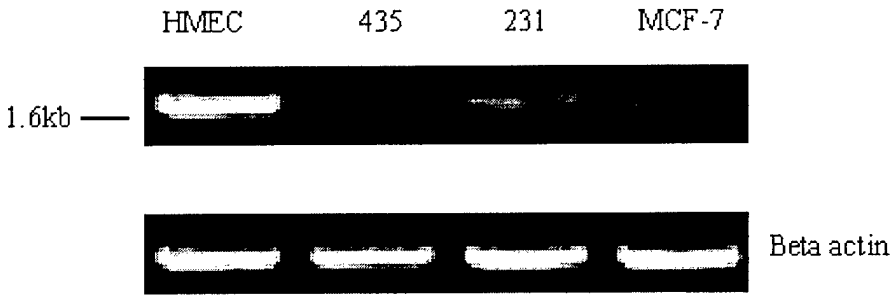
Lane1: Control sample of HMEC cell lysate
Lane 2: Soluble HMEC cell lysate + FP-probe
Lane 3: Control sample of 435 cell lysate
Lane 4: Soluble 435 cell lysate + FP-probe
Lane 5: Control sample of 231 cell lysate
Lane 6: Soluble 231 cell lysate + FP-probe
Lane 7: Control sample of MCF-7 cell lysate
Lane 8: Soluble MCF-7 cell lysate + FP-probe

To find out the identity of this protein, I immuno-precipitated the serine hydrolase-probe complex from the soluble protein lysate of HMEC cells with an antibody against red fluorescence (Tamara) and resolved the complex on 10% SDS-PAGE. The protein indicated by the arrow was excised from the gel and analyzed by mass spectrometry for protein identification. This procedure was performed three times to ensure the accuracy of the protein identity. The same protein identify was found in three independent experiments to be human carboxylesterase-2 (hCE-2).

Result 2: hCE-2 was differentially expressed in normal epithelial and breast cancer cells at the transcriptional level.

The comparison of serine hydrolases profile between the normal epithelial cells and breast cancer cells rendered the information that the catalytic activity of hCE-2 was differentially regulated in normal and tumor cells. As an initial step of verifying this finding, I performed RT-PCR to see whether the level of regulation exists at the transcriptional level. Total RNA was isolated from each cell line using Rneasy mini kit (Qiagen, CA). 2ug of RNA from each cell line was reversely transcribed and amplified using hCE-2 specific primers and Superscript one-step RT-PCR with Platinum Taq (Invitrogen, CA). RNA integrity control was set up concurrently using the β-actin primers. RT-PCR products were visualized on 1% agarose gel. The expected DNA band corresponding to the length of hCE-2 cDNA was observed in RNA sample of HMEC cells, but significantly lower in RNA from three breast cancer cell lines (figure 2).

Figure 2.



Subsequently, the gene expression of hCE-2 in HMEC cells and breast cancer cells was analyzed more quantitatively by real time PCR. hCE-2 expression was found at least 10 fold lower in breast cancer cells than HMEC cells (data not shown). Hence, I concluded that the differential regulation of hCE-2 catalytic activity seen in activity profiling could be attributed from the regulation at the transcription level.

Result #3: hCE-2 was down regulated in breast cancer cells at the protein level.

After verifying the gene regulation of hCE-2, I started to generate antibody against the protein to further verify the expression of hCE-2 at the protein level. I generated a recombinant form of hCE-2 protein N-terminus fragment, including amino acid 1-360. This recombinant form of hCE-2 has a 6Xhistine tag at the N-terminus so the molecular weight of the protein is around 16kd. Polyclonal antibody against hCE-2 was generated using the recombinant fragment of hCE-2, and total IgG was purified from the anti-serum for western blotting analysis of cell lysates. From the western blot

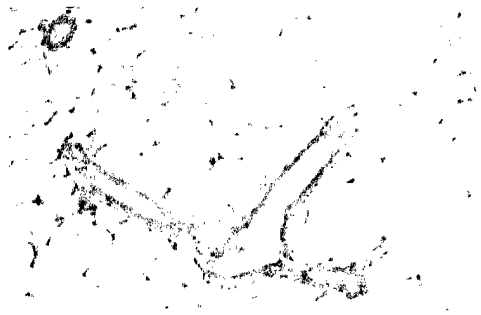
analysis, I observed that hCE-2 was down regulated at the protein level as well in breast cancer cells (figure 3).

Figure 3.

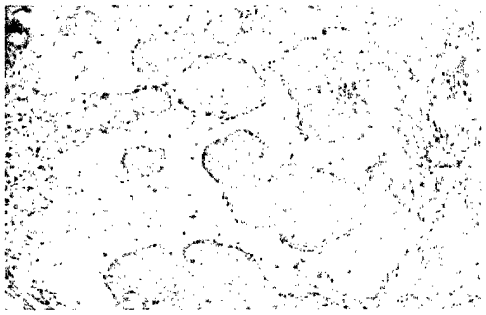
Result #4: The differential expression of hCE-2 was further verified by immunostaining of human tissue.

To further confirm my observation that hCE-2 is differentially regulated in normal tissue and in tumor, I made immuno-staining of normal human breast tissue and breast cancer tissue using my antibody (figure 4).

Figure 4.



Human breast tissue



Human infiltrating duct carcinoma

Specific staining was observed in epithelial cells inside of the duct structure of human breast tissue. However, no staining was seen in the tumor cells. This result provides further evidence that hce-2 is down regulated in tumor cells.

CONCLUSIONS AND FUTURE PLANS:

Large body of work on studying protease expression in diseases, especially in cancer progression, has been reported and continues to grow. The new technology of using activity-based probe such as FP-biotin described here will direct our attention to where the protease activity is occurred rather than its expression. This aspect of protease study is important because many proteases are synthesized as zymogens which need to be activated to perform their functions. Identifying where

the protease activity occurs will facilitate the study of the biological consequences of their function rather than the mere presence of the protein. In addition to the advantage of performing the activity-based study, the broad affinity probe such as FP-biotin can facilitate the functional analysis of these enzymes in health and disease.

In this report, I applied the affinity probe to analyze the normal epithelial cells and several breast cancer cells. Through extensive screening, I found one serine hydrolase, human carboxylesterase-2 (hCE-2), was down regulated in breast cancer cells comparing to the normal epithelial cells. The down regulation was at the level of transcription and translation as shown in the result section. Initial verifications were made by western blot analysis and immuno-staining of human tissues using the hCE-2 antibody that I generated. At present time, I also have some indication that hCE-2 might also be down regulated in prostate and colon cancer cells (data not shown). Therefore, my focus of the future will be to demonstrate the utility of hCE-2 as a diagnostic marker and study the function of hCE-2 in relation to tumor progression.

Future Plans:

1. Generate tumors by injecting 435 cells as proposed in my fellowship proposal and perform global analysis of serine hydrolase activity comparing normal mammary gland and primary tumors generated by 435 cells.
2. If new targets could be observed, identify the proteins.
3. Further characterize the hCE-2 antibody to establish the specificity profile of the antibody.
4. Study the expression pattern of hCE-2 in normal tissue by immuno-staining and commercially available tissue array.
5. Study the expression pattern of hCE-2 in other types of normal and cancer comparison such as normal human prostate vs. prostate cancer tissue.
6. Overexpress hCE-2 in 435 breast cancer cells to study the effect of hCE-2 expression in tumor cells.
7. Compare the primary tumor size and the extent of metastasis of 435 cells and hCE-2 transfected 435 cells using the orthotopic breast cancer mouse model described in the original fellowship proposal.
8. Investigate function(s) of hCE-2 by genomic approaches such as cDNA microarry and proteomic approaches such as 2D-gel electrophoresis.

KEY RESEARCH ACCOMPLISHMENTS:

- Selective MMP-14 substrate sequences were found via substrate phage display library.
- Identification of a potential tumor suppressor protein, hCE-2 in human mammary epithelial cells.

Reportable Outcomes:

One accepted manuscript are listed as the following:

Steven J. Kridel, Hisako Sawai, Boris I. Ratnikov, Emily I. Chen, Weizhong Li, Adam Godzik, Alex Y. Strongin, and Jeffrey W. Smith (2002) *Journal of Biological Chemistry* **in press.**

Conclusions/Implications:

- New families of MMP-14 substrates were found via substrate phage display library.
- Further proof of the utility of the substrate phage display library screening method developed in our lab can be useful to discriminate the catalytic activity of matrix metalloproteinases.
- A potential tumor suppressor protein was found by global analysis of normal and breast cells.

REFERENCES:

1. Basbaum, C. B., and Werb, Z. (1996) *Curr. Opin. Cell Biol.* **8**(5), 731-8
2. Werb, Z., and Yan, Y. (1998) *Science* **282,** 1279-1280
3. Nagase, H., and Woessner, J. F., Jr. (1999) *Journal of Biological Chemistry* **274**(31), 21491-4
4. Van Wart, H. E., and Birkedal-Hansen, H. (1990) *Proc Natl Acad Sci U S A* **87**(14), 5578-82
5. Nagase, H. (1997) *Biol. Chem.* **378,** 151-160
6. Kahari, V. M., and Saarialho-Kere, U. (1999) *Annals of Medicine* **31**(1), 34-45
7. Leppert, D., Ford, J., Stabler, G., Grygar, C., Lienert, C., Huber, S., Miller, K. M., Hauser, S. L., and Kappos, L. (1998) *Brain* **121**(Pt 12), 2327-34
8. Poulsom, R., Hanby, A. M., Pignatelli, M., Jeffery, R. E., Longcroft, J. M., Rogers, L., and Stamp, G. W. (1993) *Journal of Clinical Pathology* **46**(5), 429-36
9. Davies, B., Miles, D. W., Happerfield, L. C., Naylor, M. S., Bobrow, L. G., Rubens, R. D., and Balkwill, F. R. (1993) *British Journal of Cancer* **67**(5), 1126-31
10. Boag, A. H., and Young, I. D. (1994) *American Journal of Pathology* **144**(3), 585-591
11. Sier, C. F., Kubben, F. J., Ganesh, S., Heerding, M. M., Griffioen, G., Hanemaaijer, R., van Krieken, J. H., Lamers, C. B., and Verspaget, H. W. (1996) *British Journal of Cancer* **74**(3), 413-7
12. Pei, D., and Weiss, S. J. (1996) *Journal of Biological Chemistry* **271**(15), 9135-40
13. Barrett, A. J., Rawlings, N. D., and Woessner, J. F. (eds) (1998) *Handbook of Proteolytic Enzymes*, Academic Press, San Diego
14. Ohuchi, E., Imai, K., Fujii, Y., Sato, H., Seiki, M., and Okada, Y. (1997) *Journal of Biological Chemistry* **272**(4), 2446-51
15. Sato, H., Takino, T., Okada, Y., Cao, J., Shinagawa, A., Yamamoto, E., and Seiki, M. (1994) *Nature* **370**(6484), 61-5
16. Strongin, A. Y., Collier, I., Bannikov, G., Marmer, B. L., Grant, G. A., and Goldberg, G. I. (1995) *Journal of Biological Chemistry* **270**(10), 5331-8
17. Imai, K., Ohuchi, E., Aoki, T., Nomura, H., Fujii, Y., Sato, H., Seiki, M., and Okada, Y. (1996) *Cancer Research* **56**(12), 2707-10
18. Sato, H., Kinoshita, T., Takino, T., Nakayama, K., and Seiki, M. (1996) *FEBS Letters* **393**(1), 101-4
19. Deryugina, E. I., Ratnikov, B., Monosov, E., Postnova, T. I., DiScipio, R., Smith, J. W., and Strongin, A. Y. (2001) *Exp. Cell Res.* **263**(2), 209-223
20. Deryugina, E. I., Bourdon, M. A., Jungwirth, K., Smith, J. W., and Strongin, A. Y. (2000) *International Journal of Cancer* **86**(1), 15-23
21. Belkin, A. M., Akimov, S. S., Zaritskaya, L. S., Ratnikov, B. I., Deryugina, E. I., and Strongin, A. Y. (2001) *Journal of Biological Chemistry* **276**(21), 18415-22
22. Kajita, M., Itoh, Y., Chiba, T., Mori, H., Okada, A., Kinoh, H., and Seiki, M. (2001) *Journal of Cell Biology* **153**(5), 893-904
23. Deryugina, E. I., Ratnikov, B. I., Postnova, T. I., Rozanov, D. V., and Strongin, A. Y. (2001) *Journal of Biological Chemistry* **In press.**
24. Holmbeck, K., Bianco, P., Caterina, J., Yamada, S., Kromer, M., Kuznetsov, S. A., Mankani, M., Robey, P. G., Poole, A. R., Pidoux, I., Ward, J. M., and Birkedal-Hansen, H. (1999) *Cell* **99**(1), 81-92

25. Zhou, Z., Apte, S. S., Soininen, R., Cao, R., Baaklini, G. Y., Rauser, R. W., Wang, J., Cao, Y., and Tryggvason, K. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**(8), 4052-7
26. Kridel, S. J., Chen, E., Kotra, L. P., Howard, E. W., Mobashery, S., and Smith, J. W. (2001) *J. Biol. Chem.* **276**(23), 20572-20578
27. Deng, S. J., Bickett, D. M., Mitchell, J. L., Lambert, M. H., Blackburn, R. K., Carter, H. L., Neugebauer, J., Pahel, G., Weiner, M. P., and Moss, M. L. (2000) *Journal of Biological Chemistry* **40**(6), 31422-31427
28. Smith, M. M., Shi, L., and Navre, M. (1995) *J Biol Chem* **270**(12), 6440-9
29. Liu, Y., Patricelli, M. P., and Cravatt, B. F. (1999) *Proceedings of the National Academy of Sciences of the United States of America* **96**(26), 14694-9
30. Kalafatis, M., Egan, J. O., van 't Veer, C., Cawthern, K. M., and Mann, K. G. (1997) *Critical Reviews in Eukaryotic Gene Expression* **7**(3), 241-80
31. Smyth, M. J., O'Conner, M.D., Trapani, J.A. (1996) *J. Leukocyte Biol.* **60,** 555-562
32. Yoshida, S., and Shiosaka, S. (1999) *International Journal of Molecular Medicine* **3**(4), 405-9
33. Cravatt, B. F., Giang, D. K., Mayfield, S. P., Boger, D. L., Lerner, R. A., and Gilula, N. B. (1996) *Nature* **384**(6604), 83-7
34. Steiner, D. F. (1998) *Curr. Opin. Chem. Biol.* **2,** 31-39
35. Low, M. E. (1997) *J. Nutr.* **127,** 549-557
36. Mignatti, P., and Rifkin, D. B. (1996) *Enzyme & Protein* **49**(1-3), 117-37
37. Kato, G. J. (1999) *Hum. Mutat.* **13,** 87-98
38. DeClerck, Y. A., Imren, S., Montgomery, A. M., Mueller, B. M., Reisfeld, R. A., and Laug, W. E. (1997) *Advances in Experimental Medicine & Biology* **425,** 89-97
39. Khan, A. R., and James, M. N. (1998) *Protein Science* **7**(4), 815-36
40. Whisstock, J., Skinner, R., and Lesk, A. M. (1998) *Trends in Biochemical Sciences* **23**(2), 63-7
41. Roberts, R. M., Mathialagan, N., Duffy, J. Y., and Smith, G. W. (1995) *Critical Reviews in Eukaryotic Gene Expression* **5**(3-4), 385-436

APPENDICES:

One accepted manuscript is included as a supplement of the annual summary update.

JBC Papers in Press. Published on April 16, 2002 as Manuscript M111574200

A Unique Substrate Binding Mode Discriminates Membrane Type 1-Matrix Metalloproteinase (MT1-MMP) from other Matrix Metalloproteinases

Steven J. Kridel, Hisako Sawai, Boris I. Ratnikov, Emily I. Chen, Weizhong Li, Adam Godzik, Alex Y. Strongin and Jeffrey W. Smith*

The Program on Cell Adhesion
The Cancer Research Center
The Burnham Institute
10901 North Torrey Pines Road
La Jolla, CA 92037

*Corresponding author:
858-646-3121
jsmith@burnham.org

Running title: Substrate recognition by Matrix Metalloproteinases.

Summary

In our study, we characterized the substrate recognition properties of membrane type-1 matrix metalloproteinase (MT1-MMP/MMP-14) a key enzyme in tumor cell invasion and metastasis. A panel of optimal peptide substrates for MT1-MMP was identified using substrate phage display. The substrates can be segregated into four groups based on their degree of selectivity for MT1-MMP. Substrates with poor selectivity for MT1-MMP, are predominately comprised of the Pro-X-X X $_{Hy}$ motif that is recognized by a number of MMPs. Highly selective substrates lack the characteristic Pro at the P_3 position; instead they contain an Arg at the P_4 position. This P_4 Arg is essential for efficient hydrolysis and for selectivity for MT1-MMP. Molecular modeling indicates that the selective substrates adopt a linear conformation that extends along the entire catalytic pocket of MT1-MMP, while non-selective substrates are "kinked" at the conserved P_3 Pro residue. Importantly, the selective substrates can be made non-selective by insertion of a proline kink at P_3, without significantly reducing overall k_{cat}/K_m values. Altogether the study provides a structural basis for selective and non-selective substrate recognition by MT1-MMP. The findings in this report are likely to explain several aspects of MT1-MMP biology.

Introduction

Matrix metalloproteinases are a family of zinc-dependent proteases that function in a number of physiologic remodeling events including wound healing, mammary gland involution, and bone resorption (1-3). The MMPs are also linked to pathophysiologic conditions like tumor progression, inflammation and arthritis (4,5). Consequently, the MMPs have been explored aggressively as drug targets.

The MMP family can be segregated into two groups, the soluble type and the membrane-type. The membrane-type MMPs are tethered to the plasma membrane by either a transmembrane domain or a GPI linkage (6,7). Membrane Type-1 Matrix Metalloproteinase (MT1-MMP, MMP-14) is the prototypic member of the MT-MMPs, and its expression has been associated with a variety of cellular and developmental processes as well as multiple pathophysiologic conditions. Similar to most other MMPs, MT1-MMP cleaves a number of matrix proteins including collagen, fibronectin and vitronectin (8-10). Other work shows that MT1-MMP has substrates that extend beyond extracellular matrix proteins. MT1-MMP is known to convert pro-MMP-2 to active protease (6,11-13) , apparently through its tri-molecular complex with TIMP-2 and the $\alpha_v\beta_3$ integrin (14,15). Recent work shows that cleavage of cell surface molecules such as CD44, pro-α_v integrin and transglutatminase by MT1-MMP modulates cell migration (16-18).

The importance of the broad substrate recognition specificity of MT1-MMP is consistent with the complex phenotype of the MT1-MMP-deficient mice. These mice exhibit overt dwarfism, but close inspection also showed several additional defects including arthritis, delayed ossification of bone, and the inability to respond to angiogenic stimuli (19,20). Clearly, MT1-MMP is likely to have multiple physiologic and pathophysiologic substrate. Knowing the identity of these substrates, and the structural basis for their recognition, would provide an additional level of understanding of MT1-MMP. To this end, we used an unbiased substrate phage selection to identify optimal peptide substrates for MT1-MMP. Four groups of substrates, with varying degrees of selectivity were

identified. As anticipated, non-selective substrates comprised primarily of the Pro-X-X-↓-X_{Hy} sequence were identified. These substrates are collagen-like, and have emerged as a canonical, and generally non-selective, MMP recognition motif (21-23). We also identified substrates that are recognized by MT1-MMP and both gelatinases (MMP-2 and MMP-9). Most significantly, a group of highly selective substrates for MT1-MMP were identified. Each of the highly specific substrates is devoid of a Pro residue at the P_3 position and contain a critical Arg at the P_4 position. Modeling of the selective and non-selective substrates bound to the catalytic pocket of MT1-MMP indicates two separate binding modes. These modes are distinguished by the degree of distortion at the P_3 position of substrate. These observations provide a solid structural basis for the recognition of distinct classes of substrates by MT1-MMP and for the design of highly specific inhibitors of this enzyme.

Experimental procedures

Construction of substrate phage display library. Substrate phage libraries were generated using a modified version of the fUSE5 phagemid (24-26). A FLAG epitope was engineered at the N-terminus of the geneIII protein by annealing oligonucleotides 5′-CCGGGTTTGTCGTCGTCGTCTTTGTAGTCGGTAC-3′ and 5′-CGACTACAAAGACGACGACGACAAAC-3′ and ligating them into fUSE5 at the Kpn I and Xba I restriction sites. The random hexamers were generated by PCR extension of the template oligonucleotide 5′-GGGGAGGCCGACGTGGCCGTCATCAGGCGGCTCAGGC$(NNK)_6$ACGG CCTCTGGGGCCGAAAC-3′, where N is any nucleotide and K is either G or T. The template oligonucleotide also encodes an SGGSG linker positioned in between the FLAG epitope and the random hexamer. A primer oligonucleotide 5′-AATTTCTAGTTTCGGCCCCAGAGGC-3′ and the template oligonucleotide were mixed and heated at 65°C for two minutes. The heating block was switched off and allowed to cool passively to 40°C to allow annealing of the extension oligonucleotide to the template oligonucleotide. Elongation of the template oligonucleotide was performed using Sequenase (USB, Cleveland, OH) (27). The final cDNA product was precipitated with ethanol, re-suspended in water and digested with Sfi I. The DNA insert and fUSE5 were mixed and ligated at a 5:1 molar ratio and electroporated into *E. coli* MC1061(F-). Several phage were selected for sequencing to confirm randomness in the insert sequences and the correct reading frame.

Cloning and expression of recombinant MT1-MMP, MMP-2 and MMP-9 catalytic domains. Recombinant versions of the catalytic domains of MMP-2 and MMP-9 were expressed, purified and activated as described previously (21,28). The catalytic domain of MT1-MMP was amplified from a human Universal QUICK-Clone cDNA library (Clontech, Palo Alto, CA) by PCR using the following primers 5′-catatgTACGCCATCCAGGGTCTCAAATGG-3′, and 5′-gaattcttaCCCTGACTCACCCCCATAAAGTTGC-3′. The *Nde*I and *Eco*RI restriction sites and TAA stop codon are shown in lower case. The amplified PCR product was

digested with *Nde*I and *Eco*RI and cloned into the corresponding sites of pFLAG-ATS (Sigma, St. Louis, MO). The resulting plasmid (pGM-MT1) was transformed into BL21(DE3) *E. coli* and grown in 2YT supplemented with ampicillin (50 μg/ml). Protein expression was induced by the addition of IPTG to a final concentration of 0.1 mM. After three hours of induction, the bacteria were collected, lysed, and inclusion bodies containing the catalytic domain of MT1-MMP were isolated by centrifugation.

Recombinant MT1-MMP was purified from inclusion bodies by anion-exchange chromatography on a Mono Q column (Amersham-Pharmacia, Piscataway, NJ) in the presence of 8 M urea. The purified material was refolded by dilution to 100 μg/ml in 50 mM HEPES, pH 7.0, 5 mM $CaCl_2$ and 50 μM $ZnCl_2$, followed by an incubation at 12°C for one hour. The refolded MT1-MMP was immediately used in activity assays. The concentration of active enzyme was determined by active site titration. Active site titration of MT1-MMP was performed with the hydroxomate inhibitor AG3340 (K_i= $3x10^{-10}$ M) and TIMP-2. The titrations of active MT1-MMP were equivalent, thus AG3340 was used to titrate the enzyme for this study. After pre-incubation, the steady-state rate of hydrolysis of the fluorogenic substrate Mca-PLGL-Dnp-AR-NH_2 (Bachem, King of Prussia, PA) was determined at ambient temperature using an fMax fluorescence microplate reader (Molecular Devices, Sunnyvale, CA). The steady-state rate was plotted as a function of inhibitor concentration and fitted with the equation: $V = SA(E_0 - 0.5\{(E_0 + I + K_i) - [(E_0 + I + K_i)^2 - 4E_0I]^{0.5}\})$, where V is the steady-state rate of substrate hydrolysis; SA is specific activity (rate per unit enzyme concentration); E_0 is enzyme concentration; I is inhibitor concentration; and K_I is the dissociation constant of the enzyme-inhibitor complex (29).

Phage selection of MT 1-MMP substrates. An aliquot (2×10^{10} phage) of the substrate phage library was incubated with 5 μg/ml of MT1-MMP in 50 mM Tris, 7.4, 100 mM NaCl, 10 mM $CaCl_2$, 0.05% Brij-35, and 0.05% BSA for 1 hour at 37°C. A control selection was performed without protease. The cleaved phage were separated from the non-cleaved phage by immuno-depletion. The M1 anti-FLAG monoclonal antibody (100 μg) was added to the phage samples and then incubated for 18 hours with rocking at 4°C.

The phage-antibody complexes were precipitated by the addition of 100 µl Pansorbin (Calbiochem, San Diego, CA). The cleaved phage remaining in the supernatant were amplified using K91 *E. coli* as described previously (30,31), and were then used for subsequent rounds of substrate selection.

Substrate phage ELISA. Hydrolysis of individual phage substrates was measured using a modified ELISA. Wells of a 96-well microtiter plate were coated with anti-M13 antibody (Amersham-Pharmacia) at 2.5 µg/ml in PBS, overnight at 4°C. The wells were blocked for one hour at room temperature in TBS-T (50 mM Tris, pH 7.8, 150 mM NaCl, 0.2% Tween-20) containing 10 mg/ml BSA. After blocking, 150 µl of supernatant from an overnight phage culture was added to each well and incubated for 2 hours at 4°C to allow for phage capture. Unbound phage were removed with three washes of ice cold TBS-T. To assess hydrolysis, MT1-MMP (3.0 µg/ml) was added to the appropriate wells in Incubation Buffer (50 mM Tris, pH 7.4, 100 mM NaCl, 10 mM $CaCl_2$, 0.05% BSA, 0.05% Brij-35) for 2 hours at 37°C. Control wells lacked protease. The protease solution was removed and the wells were washed four times with ice cold TBS-T. To determine the extent of hydrolysis of the peptides on phage by proteinase, anti-FLAG polyclonal antibody (1.8 µg/ml in TBS-T with 1 mg/ml BSA) was added to each well and the plates were incubated at 4°C for one hour. Binding of anti-FLAG antibody to FLAG epitope was measured with an HRP-conjugated goat anti-mouse IgG antibody (BioRad, Hercules, CA) followed by detection at 490 nm. The extent of hydrolysis, taken as a measure of substrate hydrolysis, was calculated by the ratio of the O.D. at 490 nm of the protease-treated samples versus controls.

Determination of scissile bonds. The cleavage site for MT1-MMP within peptide substrates was determined using MALDI-TOF mass spectrometry. MT1-MMP (23 nM) was incubated with 100 µM of each peptide sample in 50 mM Tris, pH 7.4, 100 mM NaCl, 10 mM $CaCl_2$ for 2 hours at 37°C. Following hydrolysis, the peptide samples were prepared according to methods described previously (32,33). The mass of the cleavage products was determined using a Voyager DE-RP MALDI-TOF mass spectrometer

(PerSeptive Biosysytems, Framingham, MA). In all cases, the observed fragments corresponded to a single cleavage site.

Kinetic measurements of peptide hydrolysis. The kinetic parameters of substrate hydrolysis were measured using a fluorescamine incorporation assay that has been described previously (34). Briefly, MT1-MMP, MMP-2 and MMP-9 were incubated with individual peptide substrates at concentrations ranging from 100-800 μM in 50 mM Tris, pH 7.5, 100 mM NaCl, 10 mM $CaCl_2$, 50 μM $ZnCl_2$. At selected time points the reactions were stopped by the addition of 1,10-phenanthroline. Peptide hydrolysis was determined by the addition of fluorescamine, followed by detection at λ_{ex} 355 nm and λ_{em} 460 nm. The data were transformed to double reciprocal plots (1/[S] vs $1/v_i$) to determine K_m and k_{cat} (34). Similar results were obtained using different batches of protease. For some substrates, K_m and k_{cat} could not be determined individually, but the specificity constant, k_{cat}/K_m, was derived by the equation: $k_{cat}/K_m = v_i/(E_0)(S_0)$ (35).

Modeling of substrate peptides with MT1-MMP. The complexes of MT1-MMP and cognate substrates were modeled using the Sybyl software package available from Tripos (http://www.tripos.com) using the MT1-MMP spatial structure taken from PDB file 1bqq (36-38).

Results

Selection of peptide substrates for MT1-MMP.

The catalytic domain of MT1-MMP was used in three successive rounds of selection to isolate optimum substrates from a phage library of 2.4×10^8 independent clones. Two-hundred clones were screened for their ability to be hydrolyzed by MT1-MMP using the substrate phage ELISA we previously reported (21,28). Forty of the two-hundred clones were efficiently cleaved by MT1-MMP and were carried forward for additional analysis.

In a second round of analysis, substrate phage ELISAs were used to compare relative rates of hydrolysis of phage peptide substrate by MT1-MMP, MMP-2 and MMP-9 catalytic domains. Based on the activity profiles resulting form this analysis, the forty phage substrates could be segregated into four distinct groups: *i*) those recognized by all three MMPS, *ii*) those recognized equally by MT1-MMP and MMP-2, *iii*) those recognized equally by MT1-MMP and MMP-9, and *iv*) those selective for MT1-MMP (Table I).

The amino acid sequence of each phage peptide was determined by sequencing the corresponding DNA inserts of the selected phage, and the peptides are presented in Table I. These sequences reveal the structural features common to each group of substrates. The non-selective substrates of group I all contain Pro residues. Interestingly, all of the Pro residues were expressed at the same site in the random peptides. Upon first inspection, these substrates appear similar to the Pro-X-X-↓-Hy motif that is a commonly cleaved by many MMPs (21-23). This motif represents a canonical collagen-like MMP recognition motif (9). The substrates in group II, which are susceptible to cleavage by MT1-MMP and MMP-2, were primarily comprised of substrates that appeared to be similar in primary sequence to an MMP-2 recognition motif we recently identified (39). The substrates in group III, which are susceptible to cleavage by MT1-MMP and MMP-9, were relatively small in number, however, consequently no consensus could be defined. The substrates that were highly selective for MT1-MMP are in group IV and appeared to give little initial indication of consensus. It is noteworthy that Pro is absent from these peptides, however. Given the contrast in substrates selectivity between

groups I and IV, we focused on trying to determine the structural basis for their distinct reactivity with MT1-MMP.

Identification of scissile bonds within MT1-MMP peptide substrates.

Although the position of the scissile bond in the phage substrates could generally be surmised by the position of hydrophobic residues (Leu/Ile) that normally occupy the deep $S_{1'}$ pocket, we confirmed the placements by analyzing the hydrolysis of synthetic peptides. Peptides were synthesized to mimic the peptides displayed by the individual phage clones. Following exposure to MT1-MMP catalytic domain, the masses of the cleavage products were determined with MALDI-TOF mass spectrometry (data not shown). The MALDI analysis confirmed that hydrophobic amino acids occupied the $P_{1'}$ position of the MT1-MMP substrates (Table II). In addition, a polar amino acid appears to be preferred at the $P_{2'}$ position. Furthermore, an alignment of the group IV peptides based on their scissile bonds demonstrates a preference for Arg at the P_4 position. Two substrates, A20 and A167, from group IV that lacked Arg at P_4 were also examined. The masses of cleavage products of these two peptides indicate that the P_4 position is Phe or Lys, respectively, which both contain long side chains like Arg that can participate in hydrophobic interactions (see below).

Kinetic analysis of substrate hydrolysis by MT1-MMP.

Results from the substrate phage ELISA (Table I) indicate a rank order of substrate preference within each group of MT1-MMP substrates. To precisely quantify the rate of hydrolysis of the peptide substrates in group IV by MT1-MMP, the specificity constant (k_{cat}/K_m) of selected peptides was determined (Table II). In these studies only MMP-9 was used for comparisons with MT1-MMP because MMP-9 and MMP-2 share similar substrate specificity and cleavage efficiency. Each proteinase was incubated with increasing concentrations of peptide. Peptide hydrolysis was measured by incorporation of fluorescamine onto newly formed amino-termini. From these measurements, k_{cat} and K_m values were derived for each peptide using double reciprocal plots of 1/[S] *versus* $1/v_i$. Among the peptides, overall k_{cat}/K_m values ranged from 8,200 $M^{-1}s^{-1}$ to 777,200 $M^{-1}s^{-1}$, and the selectivity of these peptides ranged from 2.2-fold to 83-fold when compared

to hydrolysis by MMP-9. Since nearly one-half of the MT1-MMP selective substrates lack Arg at P_4, there may be other modes of selective recognition. Nevertheless, we suspect that some of these other peptides are related to the Arg containing peptides. For example, Peptides A167 and A20, which lack Arg at P_4 showed the lowest k_{cat}/K_m values, and the lowest selectivity ratios. Yet, Phe and Lys, which occupy P4 in these peptides, can participate in similar hydrophobic interactions.

Assessing the role of Arg at P_4 in substrate recognition by MT1-MMP.

Because of the preponderance of Arg at the P_4 position among the group IV substrates, we sought to determine the requirement for this residue in hydrolysis by MT1-MMP. Three peptides that contain this P_4 Arg, A42, B175 and A176 were selected for mutational analysis. The P_4 position in each synthetic peptide was changed from Arg to Ala and k_{cat}/K_m values were measured (Table II). This substitution decreased k_{cat}/K_m of peptide A176 from 49,100 $M^{-1}s^{-1}$ to 5,900 $M^{-1}s^{-1}$. Similarly the same substitution in B175 reduced k_{cat}/K_m from 63,600 $M^{-1}s^{-1}$ to 10,400 $M^{-1}s^{-1}$. The Arg to Ala substitution had a less significant effect in the A42 peptide (three-fold reduction). We suspect that the very high k_{cat}/K_m of this peptide (777,200 $M^{-1}s^{-1}$) indicates the presence of several additional contacts that favor substrate recognition by MT1-MMP. Consequently, substitution of the Arg at P_4 has less of an effect on A42. Altogether, the mutational analyses support our hypothesis that Arg at P_4 is a key contact residue in the group IV substrates

Making selective substrates non-selective.

An obvious distinction between the selective and non-selective substrates is the presence of Pro at the P_3 position. We sought to determine if the presence of a Pro at this position of the group IV substrates would be sufficient to override their selective docking. For this purpose, we chose the peptide with the highest selectivity for MT1-MMP (A176). When the P_3 Ser of A176 was replaced with Pro, the k_{cat}/K_m by MT1-MMP was reduced by only 32%. Conversely, the hydrolysis of this peptide by MMP-9, a measure of selectivity, was increased 58-fold. Hence, substitution of Pro at the P_3 position converted a substrate selective for MT1-MMP to a substrate that is recognized equally

well by both MMPs. A similar substitution was made into the A176A peptide, in which the Arg at P_4 of A176 was changed to Ala. This substitution caused a 9-fold reduction in k_{cat}/K_m ratio for MT1-MMP compared to the parent peptide. Yet, when Pro was substituted into the P_3 position of this peptide the k_{cat}/K_m ratio for MT1-MMP increased by approximately six-fold. Consequently, the deleterious effect of the Arg to Ala substitution at P_4 was rescued by substitution of Pro at P_3. Equally as significant, this substitution reversed the selectivity ratio to favor MMP-9. These findings support the idea that MT1-MMP recognizes substrates in two distinct modes. One mode, demonstrated by the group I peptides, makes use of the P_3 and P_1 positions as dominant contact points. The other mode, demonstrated by the selective group IV peptides, appears to primarily use the P_4 and P_1 subsites as key contacts.

Modeling substrates in the catalytic pocket of MT1-MMP.

To visualize the differences in these two binding modes, molecular modeling studies were conducted. A 3-D model of the catalytic domain of MT1-MMP was generated from the coordinates of its known crystal structure (36). The catalytic domain was then docked with peptides A176 (Ac-SGRSEN↓IRTA-NH_2), and A176P (Ac-SGRPEN↓IRTA-NH_2) using Sybyl. Because interactions between MMP and substrate at the $S_{1'}$ subsite are known to make a major contribution to substrate binding, interactions around this site were kept constant to constrain the modeling procedure. Two restrictions were applied. First, the hydrogen bond between the NH group of the $P_{1'}$ Ile residue and the carbonyl oxygen of Ala^{200} in MT1-MMP was set as a constant. Second, the interaction between side-chain of the $P_{1'}$ Ile residue of substrate and the hydrophobic patch made up of Val^{236} and the surrounding atoms in MT1-MMP was held constant. The initial docking conformation of the substrate was essentially a β-strand running anti-parallel to the neighboring β-strand in MT1-MMP (Ala^{200} to Ala^{202}). Conformation "searches" of the N-terminal and C-terminal halves of the substrate were run separately. To reduce the computation, the conformations were searched in torsion angle space, in which bond length and bond angles of the whole substrate remained fixed. These computational searches led to the hypothetical docking modes illustrated in Figure 1.

Interestingly, peptide A176 fits in the catalytic cleft of MT1-MMP with very little deviation from linearity (panel A). The Arg at the P_4 position in this peptide is slightly distorted, apparently facilitating hydrophobic contact with Gly^{116} and Phe^{204} of MT1-MMP. It would appear that these hydrophobic contacts are the dominant interactions within the S_4 pocket of MT1-MMP. In contrast to the extended linear conformation of peptide A176, a variant of this peptide, with Pro at P_3, bound with a pronounced kinked conformation. The effect of this deviation from linearity is that the P_4 Arg is not involved in hydrophobic contacts described for peptide A176. Instead, the presence of Pro at the P_3 position changes contacts such that the P_3 Pro residue interacts with Phe^{204} of MT1-MMP, in effect suppressing the ability of the P_4 Arg to make its contacts. These models illustrate the likely structural distinctions between the selective and non-selective substrate binding modes of MT1-MMP, thereby representing the apparent distinction between collagen and non-collagen substrates by MMPs.

Discussion.

The absolute level of MMP activity is regulated at three key steps. First, virtually all MMPs are regulated at the level of transcription (1,2). Second, inactive MMP zymogens are activated by a proteolytic removal of a segment of the pro-domain. Third, their natural inhibitors, the tissue inhibitors of metalloproteinases (TIMPs), can inhibit the activated proteinases. Another key aspect of MMP biology is the control at the level of substrate recognition. Substrate recognition can be strongly modulated by domains ancillary to the catalytic cleft, like the hemopexin domain, or the type II fibronectin repeats in gelatinases (40-42). These domains guide proteolysis by binding to specific substrates. The control of substrate hydrolysis at the catalytic cleft is generally thought to play a lesser role. This perception is partly driven by the overall structural similarity of the catalytic clefts of MMPs which all contain a defining zinc ion, zinc coordinating His residues, a catalytic glutamic acid, and rather deep, S_1 pocket (43,44). Indeed, it is the S_1 pocket that has been exploited to design small molecule antagonists of MMPs. The fact that many of these antagonists have a broad inhibition spectrum also supports the contention that substrate recognition is similar at the catalytic cleft. Furthermore, many MMPs recognize collagen, and collagen-like peptides, like those containing the P-X-X-X_{Hy} motif described here. Despite these indications of similarity, however, we found that even the closely related gelatinases, MMP-2 and MMP-9, show distinct recognition profiles for small peptide substrates. Others have shown that MMP-3, MMP-7 and MMP-13 also has a unique substrate recognition capability (22,23). These findings lead to questions about how encompassing, and how distinct, the substrate recognition profiles of the MMP family may be.

The objective of this study was to define the scope of substrate recognition by MT1-MMP, and to compare its recognition patterns to that of MMP-2 and MMP-9. We were particularly interested in knowing whether MT1-MMP can recognize substrates that other MMPs cannot. Using substrate phage display, we were able to identify four distinct groups of substrates with the catalytic domain of MT1-MMP. While three of the groups did not exhibit any appreciable selectivity, one of the substrate groups (group IV) is

highly selective for MT1-MMP. The defining feature of these substrates is the presence of an Arg residue at the P_4 position. The P_4 Arg residue makes major contributions to the k_{cat}/K_m ratio and to the selectivity of the substrates for MT1-MMP (Table II). To our knowledge, this is the first description of a key docking point to the left of the S_3 subsite for MMP substrates. The selective substrates have two other common features. First, they have a hydrophobic residue (often Leu or Ile) at the $P_{1'}$ position. The occupation of this position by hydrophobic residues is not unexpected because all MMPs contain a deep S_1 pocket that has a dominant role in substrate recognition. Second, the selective substrates lack Pro at the P_3 position. The lack of this residue is striking because the vast majority of MMP peptide substrates selected from this phage library contain such Pro residues (21,28).

The substrates for MT1-MMP illustrate an emerging picture of substrate recognition at the catalytic pocket. The nature of the substrates identified here strongly indicates that cooperativity between subsites is key to recognition and hydrolysis. That is, the binding of substrate at one subsite influences the binding of substrate at another. Both types of substrates contain hydrophobic residues that fit into the deep S_1 pocket of MT1-MMP. Beyond this contact point though, the two sets of substrates access different subsites within MT1-MMP. Our studies show that the P_4 Arg residue of the Group IV substrates has a key role in recognition. Molecular modeling indicates a significant role for the methylene groups of this Arg side chain in making hydrophobic contacts with Gly^{116} and Phe^{204} of MT1-MMP (see Figure 1). In contrast, the non-selective substrates of Group I primarily use the P_3 subsite as key contact point. In agreement, the insertion of a Pro into the P_3 position of the selective substrates can completely override the effects of the Arg at P_4, presumably by causing a kink in the peptide bond that disallows favorable docking of the P_4 residue.

The information on the structure of the substrates in Group IV provides a structural basis for the recognition of at least one know protein substrate for MT1-MMP. Recent work from one of our groups shows that MT1-MMP will cleave transglutaminase (16). The cleavage site within transglutaminase contains the two key features of the Group IV

substrates, an Arg at P_4 and a hydrophobic residue at P_1 . Moreover, there is no Pro residue at P_3. It is interesting to speculate that other novel substrates for MT1-MMP could be identified through a combination of structural data from Group IV substrate and genomic databases.

Although we focused primarily on the contrast between the Group I and IV substrates in this study, it should also be emphasized that substrate recognition by MT1-MMP is not limited to just these two binding modes. The substrate phage selections show that MT1-MMP recognizes at least two other sets of substrates (Table 1). The group II substrates, which are recognized by MT1-MMP and MMP-2, provide additional support for the idea that cooperative interaction among subsites is important. The consensus sequence of this family indicates that the P_3 and P_i subsites are dominant. Interestingly, most of the peptides substrates that been selected for the MMPs appear to have key contact points on both sides of the scissile bond. Given that no active MMP has been crystallized with bound substrate, attempts to co-crystallize active site mutants of MMPs with the selective substrates are certainly warranted.

The nature of the substrates identified here suggests unexpected relationships among the MMPs. The substrates selected by MT1-MMP are often recognized by MMP-2 or MMP-9. The Venn diagram in Figure 2 can describe the nature of the substrates. For example, Substrates in group I are recognized equally well by all three MMPs. Substrates in group II are recognized equally well by MT1-MMP and MMP-2. Yet, these peptides are not cleaved efficiently by MMP-9. In contrast, substrates in group III are recognized by MT1-MMP and MMP-9, but not by MMP-2. This outline of substrate recognition could ultimately be expanded to include the entire MMP family. Such an expanded view might be better described by the clustering algorithms popularized for gene array analysis (45,46), rather than Venn diagrams. It is conceivable that matching such clusters with structural information on the subsites of each MMP might give clues to detailed structural basis for the overlap and distinction in substrate recognition.

The potential for distinct methods of recognizing substrate within the catalytic cleft suggests an interesting hypothesis about another potential level of regulation of MMP activity; substrate switching. While we have investigated active site interactions in this report, the influence on substrate recognition by exosite interactions remains unknown. Protein-protein interactions at domains beyond the catalytic cleft, like the hemopexin domain, could allosterically influence the structure of the catalytic pocket. One might hypothesize a similar role for the cytoplasmic tail of MT1-MMP, which is known to localize MT1-MMP to invasive fronts and regulate trafficking and internalization (47). Inside-out signaling is a well known phenomena in the integrin field (48-50), and could be at play in regulation substrate hydrolysis by the membrane-type MMPs. The membrane type MMPs also contain an additional loop in the catalytic domain, called the MT-specific loop. Recent work shows that this loop regulates MT1-MMPs ability to activate MMP-2 (51). Although mutations within the MT-loop were without effect on peptide substrates with the P-X-X-X_{Hy} motif, the effect of such mutations on the substrate-binding mode represented by the group IV peptides is worthy of exploration. Finally, it is tempting to hypothesize that a single substrate could dynamically switch between the two binding modes during catalysis. Such a mechanism could be invoked in the hydrolysis of collagen, which is "unwound" during hydrolysis by the MMPs (52).

The results of the present study also suggest a path for defining the biological significance of each of the substrate binding modes for MT1-MMP *in vivo*. We anticipate that point mutations can be made to selectively ablate individual binding modes, without significantly affecting the other modes. The hypothesized residues that form hydrophobic contacts with the P_4 Arg side chain (group IV substrates), or that disallow the kink imposed by Pro at P_3 (group I substrates) are good initial targets for mutation. If such mutations can be made without altering recognition of other classes of substrates, the biological role of each binding mode can be determined. Consequently, future work will aim to determine if the wide-ranging phenotype of the MT1-MMP-deficient mice can be segregated into effects attributable to each individual substrate-binding mode.

Acknowledgements

Grants AR 42750 and CA69306 from the NIH, and grant 5JB0033 from the California Breast Cancer Research Program to JWS supported this work. Additional support was provided by grants CA83017 and CA77471 from the NIH, 5JB0094 form the California Breast Cancer Research Program, and grant 9849 from the Susan G. Komen Breast Cancer Foundation, all to AYS. SJK was supported by fellowships AR08505 from the National Institutes of Health and 2PD0182 from the California Cancer Research Program.

References

1. Shapiro, S. D. (1998) *Curr. Opin. Cell Biol.* **10,** 602-608
2. Basbaum, C. B., and Werb, Z. (1996) *Curr. Opin. Cell Biol.* **8**(5), 731-8
3. Werb, Z., and Yan, Y. (1998) *Science* **282,** 1279-1280
4. Seiki, M. (1999) *APMIS* **107**(1), 137-43
5. Polette, M., and Birembaut, P. (1998) *International Journal of Biochemistry & Cell Biology* **30**(11), 1195-202
6. Sato, H., Takino, T., Okada, Y., Cao, J., Shinagawa, A., Yamamoto, E., and Seiki, M. (1994) *Nature* **370**(6484), 61-5
7. Kojima, S., Itoh, Y., Matsumoto, S., Masuho, Y., and Seiki, M. (2000) *FEBS Letters* **480**(2-3), 142-6
8. Pei, D., and Weiss, S. J. (1996) *Journal of Biological Chemistry* **271**(15), 9135-40
9. Barrett, A. J., Rawlings, N. D., and Woessner, J. F. (eds) (1998) *Handbook of Proteolytic Enzymes*, Academic Press, San Diego
10. Ohuchi, E., Imai, K., Fujii, Y., Sato, H., Seiki, M., and Okada, Y. (1997) *Journal of Biological Chemistry* **272**(4), 2446-51
11. Strongin, A. Y., Collier, I., Bannikov, G., Marmer, B. L., Grant, G. A., and Goldberg, G. I. (1995) *Journal of Biological Chemistry* **270**(10), 5331-8
12. Imai, K., Ohuchi, E., Aoki, T., Nomura, H., Fujii, Y., Sato, H., Seiki, M., and Okada, Y. (1996) *Cancer Research* **56**(12), 2707-10
13. Sato, H., Kinoshita, T., Takino, T., Nakayama, K., and Seiki, M. (1996) *FEBS Letters* **393**(1), 101-4
14. Deryugina, E. I., Ratnikov, B., Monosov, E., Postnova, T. I., DiScipio, R., Smith, J. W., and Strongin, A. Y. (2001) *Exp. Cell Res.* **263**(2), 209-223
15. Deryugina, E. I., Bourdon, M. A., Jungwirth, K., Smith, J. W., and Strongin, A. Y. (2000) *International Journal of Cancer* **86**(1), 15-23
16. Belkin, A. M., Akimov, S. S., Zaritskaya, L. S., Ratnikov, B. I., Deryugina, E. I., and Strongin, A. Y. (2001) *Journal of Biological Chemistry* **276**(21), 18415-22
17. Kajita, M., Itoh, Y., Chiba, T., Mori, H., Okada, A., Kinoh, H., and Seiki, M. (2001) *Journal of Cell Biology* **153**(5), 893-904

18. Deryugina, E. I., Ratnikov, B. I., Postnova, T. I., Rozanov, D. V., and Strongin, A. Y. (2001) *Journal of Biological Chemistry* **In press.**

19. Holmbeck, K., Bianco, P., Caterina, J., Yamada, S., Kromer, M., Kuznetsov, S. A., Mankani, M., Robey, P. G., Poole, A. R., Pidoux, I., Ward, J. M., and Birkedal-Hansen, H. (1999) *Cell* **99**(1), 81-92

20. Zhou, Z., Apte, S. S., Soininen, R., Cao, R., Baaklini, G. Y., Rauser, R. W., Wang, J., Cao, Y., and Tryggvason, K. (2000) *Proceedings of the National Academy of Sciences of the United States of America* **97**(8), 4052-7

21. Kridel, S. J., Chen, E., Kotra, L. P., Howard, E. W., Mobashery, S., and Smith, J. W. (2001) *J. Biol. Chem.* **276**(23), 20572-20578

22. Deng, S. J., Bickett, D. M., Mitchell, J. L., Lambert, M. H., Blackburn, R. K., Carter, H. L., Neugebauer, J., Pahel, G., Weiner, M. P., and Moss, M. L. (2000) *Journal of Biological Chemistry* **40**(6), 31422-31427

23. Smith, M. M., Shi, L., and Navre, M. (1995) *J Biol Chem* **270**(12), 6440-9

24. Smith, G. P. (1985) *Science* **228,** 1315-1317

25. Smith, G. P. (1993) *Gene* **128,** 1-2

26. Scott, J. K., and Smith, G. P. (1990) *Science* **249**(4967), 386-90

27. Koivunen, E., Arap, W., Valtanen, H., Rainisalo, A., Medina, O. P., Heikkila, P., Kantor, C., Gahmberg, C. G., Salo, T., Konttinen, Y. T., Sorsa, t., Rouslahti, E., and Pasqualini, R. (1999) *Nature Biotechnology* **17,** August 1999

28. Kridel, S. J., Chen, E., and Smith, J.W. (2001) *Analytical Biochemistry* **In Press**

29. Knight, C. G. (1995) *Methods in Enzymology* (Barrett, A. J., Ed.), 248, Academic Press, San Diego

30. Ding, L., Coombs, G. S., Strandberg, L., Navre, M., Corey, D. R., and Madison, E. L. (1995) *Proc.Natl.Acad.Sci.U.S.A.* **92,** 7627-7631

31. Madison, E. L., Coombs, G. S., and Corey, D. R. (1995) *J. Biol. Chem.* **270,** 7558-7562

32. Landry, F., Lombardo, C. R., and Smith, J. W. (2000) *Anal. Biochem.* **279,** 1-8

33. Vorm, O., Roepstorff, P., and Mann, M. (1994) *Anal. Chem.* **66,** 3281-3287

34. Fields, G. B., Van Wart, H. E., and Birkedal-Hansen, H. (1987) *Journal of Biological Chemistry* **262**(13), 6221-6

35. Netzel-Arnett, S., Sang, Q. X., Moore, W. G., Navre, M., Birkedal-Hansen, H., and Van Wart, H. E. (1993) *Biochemistry* **32**(25), 6427-32

36. Fernandez-Catalan, C., Bode, W., Huber, R., Turk, D., Calvete, J. J., Lichte, A., Tschesche, H., and Maskos, K. (1998) *EMBO Journal* **17**(17), 5238-48

37. Humphrey, W., Dalke, A., and K, S. (1996) *J MOl Graph* **14,** 33-38

38. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Engineering* **8,** 127-134

39. Chen, E. I., Kridel, S. J., Howard, E. W., Li, W., Godzik, A., and Smith, J. W. (2002) *Journal of Biological Chemistry* **In press.**

40. Murphy, G., Willenbrock, F., Ward, R. V., Cockett, M. I., Eaton, D., and Docherty, A. J. (1992) *Biochemical Journal* **283**(Pt 3), 637-41

41. Ward, R. V., Atkinson, S. J., Reynolds, J. J., and Murphy, G. (1994) *Biochemical Journal* **304**(Pt 1), 263-9

42. O'Connell, J. P., Willenbrock, F., Docherty, A. J., Eaton, D., and Murphy, G. (1994) *Journal of Biological Chemistry* **269**(21), 14967-73

43. Stocker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Ruth, F. X., McKay, D. B., and Bode, W. (1995) *Protein Sci* **4**(5), 823-40

44. Bode, W., Fernandez-Catalan, C., Tschesche, H., Grams, F., Nagase, H., and Maskos, K. (1999) *Cellular & Molecular Life Sciences* **55**(4), 639-52

45. Ross, D. T., Scherf, U., Eisen, M. B., Perou, C. M., Rees, C., Spellman, P., Iyer, V., Jeffrey, S. S., Van de Rijn, M., Waltham, M., Pergamenschikov, A., Lee, J. C., Lashkari, D., Shalon, D., Myers, T. G., Weinstein, J. N., Botstein, D., and Brown, P. O. (2000) *Nature Genetics* **24**(3), 227-35

46. Yang, G. P., Ross, D. T., Kuang, W. W., Brown, P. O., and Weigel, R. J. (1999) *Nucleic Acids Research* **27**(6), 1517-23

47. Lehti, K., Valtanen, H., Wickstrom, S., Lohi, J., and Keski-Oja, J. (2000) *Journal of Biological Chemistry* **275**(20), 15006-13

48. Ruoslahti, E. (1996) *Tumour.Biol.* **17,** 117-124

49. Short, S. M., Talbott, G. A., and Juliano, R. L. (1998) *Mol. Biol. Cell* **9,** 1969-1980

50. Howe, A., Aplin, A. E., Alahari, S. K., and Juliano, R. L. (1998) *Current Opinion in Cell Biology* **10**(2), 220-31

51. English, W. R., Holtz, B., Vogt, G., Knauper, V., and Murphy, G. (2001) *Journal of Biological Chemistry* **276**(45), 42018-42026

52. Lauer-Fields, J. L., Tuzinski, K. A., Shimokawa, K., Nagase, H., and Fields, G. B. (2000) *Journal of Biological Chemistry* **275**(18), 13282-13290

Figure Legends

Figure 1. Three-dimensional models of distinct binding modes between substrate and MT1-MMP. To illustrate the difference in binding modes between different substrate groups into the active site of MT1-MMP, models of two representative peptides binding to MT1-MMP were constructed. The energy minimized complexes of peptides A176 (SGRSENIRTA, *panel A*) and A176P (SGRPENIRTA, *panel B*) in the catalytic domain of MT1-MMP are shown. The active site of the proteinase is shown in green, and the substrate peptide is in blue. The catalytic zinc atom is represented by an orange sphere.

Figure 2. Venn diagram of MMP substrates. Individual circles represent hypothetical substrates for MT1-MMP, MMP-2 and MMP-9. The overlapping regions imply an overlap in substrate selectivity. The MT1-MMP substrates identified in this study are represented by roman numerals corresponding to group numbers and are placed according to their defined activity profiles.

Table I

Peptide sequence and grouping of phage selected by MT1-MMP

Individual phage were analyzed after three rounds of selection by MT1-MMP. The substrates are grouped according to their ability to be hydrolyzed by the three proteinases, relative to non-treated controls, as measured by substrate phage ELISA and expressed as %Hydrolysis.

								%Hydrolysis		
	Clone #	**Peptide Sequence**						**MT1**	**MP2**	**MP9**
Group I	A79	W	L	P	S	S	I	96	91	85
	A71	R	L	P	L	G	I	94	81	95
	B32	Q	R	P	H	A	I	93	89	89
	B133	K	G	P	I	H	L	91	87	89
	B117	F	L	P	Q	G	W	90	84	58
	B95	Y	A	P	Y	G	L	88	87	85
	A101	L	D	P	W	S	Y	81	60	85
Group II	B144	L	L	G	L	R	N	91	81	21
	B164	Q	L	L	G	F	L	90	87	7
	B51	A	V	V	Q	H	L	89	88	7
	B115	N	A	L	A	S	I	88	86	0
	A112	A	R	L	G	H	F	85	91	5
	B109	G	A	V	K	H	L	81	76	15
	B145	G	A	A	R	H	I	79	72	26
	A108	N	L	R	S	K	L	79	87	3
	B173	N	L	K	S	R	V	77	85	0
	A82	D	R	Q	M	A	L	69	46	2
	A162	F	S	M	F	S	L	66	48	8
	A120	W	A	K	S	N	L	48	80	0
	A74	V	S	L	K	N	L	45	81	0
	B112	T	V	G	W	I	S	33	85	12
Group III	A56	V	A	S	G	I	Y	86	19	47
	A138	Y	A	P	Y	G	L	82	0	37
	A76	P	Q	H	V	R	Q	71	2	38
	A140	A	V	Y	F	H	M	66	0	21
	A51	L	R	A	M	Q	T	65	0	37
	A133	G	A	N	S	L	K	63	0	43
Group IV	A42	R	I	G	F	L	R	88	30	18
	A167	V	F	S	I	P	L	79	25	9
	B175	R	A	M	H	M	Y	79	6	12
	A176	R	S	E	N	I	R	79	4	2
	A20	I	K	Y	H	S	L	77	26	14
	B96	L	I	S	H	S	I	73	14	0
	B149	A	R	Y	R	W	L	72	20	3
	A156	G	V	S	M	P	F	67	11	0
	B91	V	M	S	I	R	I	61	0	0
	A69	R	S	Y	A	I	L	55	0	0
	A125	S	R	R	L	H	L	50	7	14
	A145	F	I	A	N	P	V	42	0	0

Table II

Kinetic analysis of MT1-MMP peptide substrates.

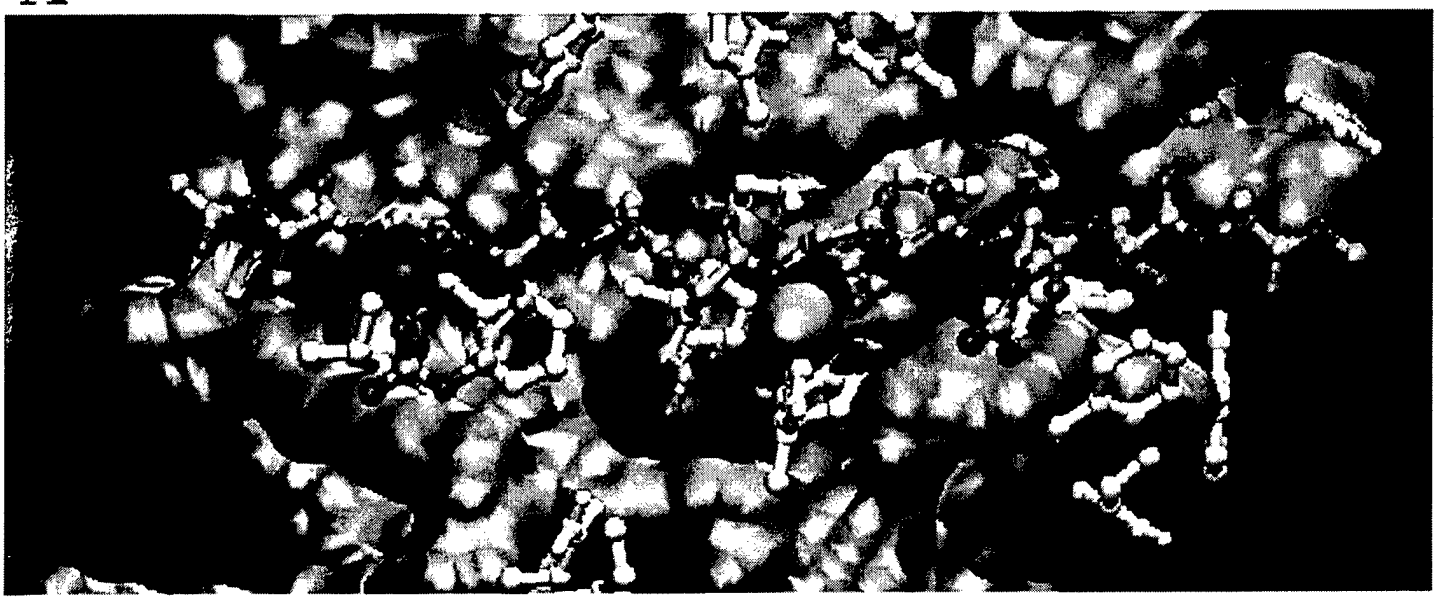
Each peptide was incubated with proteinase and hydrolysis was detected as described under "Experimental procedures". The peptide names correspond to their phage clone designation. Amino acid changes of the variant peptides are indicated in bold. Scissile bonds are indicated by ↓. Each experiment was performed in triplicate and standard deviations were 15% or below. n.a. is not able to detect.

Peptide																	k_{cat}/K_m ($M^{-1}s^{-1}$)	
																	MT1-MMP	**MMP-9**
A42			S	G	R	I	G	F	↓	L	R	T	A				777,200	20,000
A42A			S	G	**A**	I	G	F	↓	L	R	T	A				206,300	13,900
B175			S	G	R	A	M	H	↓	M	Y	T	A				63,600	9,400
B175A			S	G	**A**	A	M	H	↓	M	Y	T	A				10,400	n.a.
A176			S	G	R	S	E	N	↓	I	R	T	A				49,600	600
A176A			S	G	**A**	S	E	N	↓	I	R	T	A				5,900	400
A176P			S	G	R	**P**	E	N	↓	I	R	T	A				33,500	35,300
A176AP			S	G	**A**	**P**	E	N	↓	I	R	T	A				27,600	31,700
B149	S	G	A	R	Y	R	W	↓	L	T	A						27,500	11,600
B96	S	G	L	I	S	H	S	↓	I	T	A						15,500	7,200
A108					S	G	N	↓	L	R	S	K	L	T	A		13,900	1,600
A167	S	G	V	F	S	I	P	↓	L	T	A						12,400	5,400
A20	S	G	I	K	Y	H	S	↓	L	T	A						8,200	2,500

Figure 1

A

A176



B

A176P

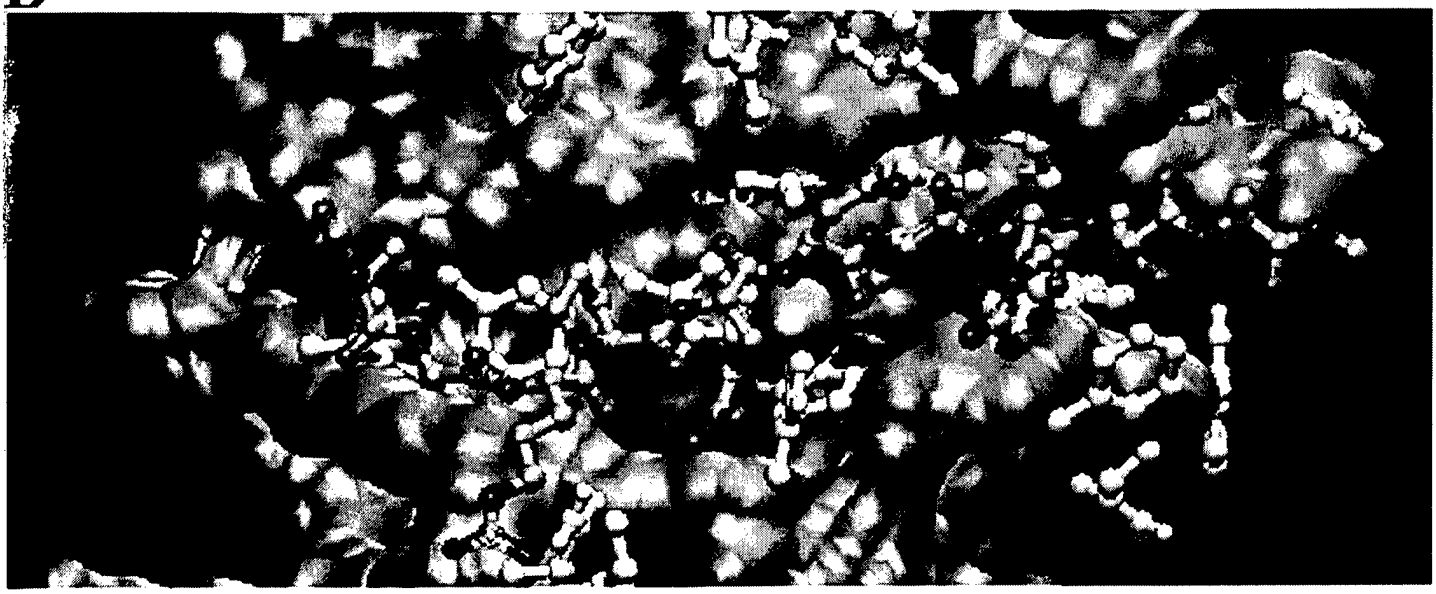


Figure 2